

THE SATISFIABILITY THRESHOLD FOR k -XORSAT

BORIS PITTEL AND GREGORY B. SORKIN[†]

ABSTRACT. We consider an “unconstrained” random k -XORSAT problem, which is a uniformly random system of m linear non-homogeneous Boolean equations over n variables, each equation containing $k \geq 3$ variables, and a “constrained” model where every variable appears in at least two equations. For $k = 3$, Dubois and Mandler proved that $m/n = 1$ is a sharp threshold for almost certain solvability of constrained k -XORSAT, and analyzed the 2-core of a random 3-regular hypergraph to extend this result to the threshold value of m/n for unconstrained 3-XORSAT. In this paper we show that $m/n = 1$ remains a sharp threshold for solvability for every $k \geq 3$, and we use Molloy’s analysis of the 2-core of a random k -regular hypergraph to extend this result to the threshold value of m/n for unconstrained k -XORSAT.

1. INTRODUCTION

An instance of k -XORSAT is given by a set of linear equations modulo 2 in which each equation consists of k variables drawn from a set of n variables, and a right hand side, which is either 0 or 1. Equivalently, it is a linear system $Ax = b$ modulo 2 in which A is an $m \times n$ 0–1 matrix each of whose row sums is k , and b is an arbitrary 0–1 vector.

Random instances of many problems of this sort undergo phase transitions around some critical ratio $\lim_{n,m \rightarrow \infty} m/n = c^*$, meaning that for any $c < c^*$, a sequence of random instances $F_{n,m}$ with $m/n \leq c$, the probability that $F_{n,m}$ is satisfiable (or possesses some similar property) approaches 1, while the probability approaches 0 if $m/n \geq c > c^*$. Friedgut [11] proved that a wide range of problems have such sharp thresholds, but with the possibility that the threshold $c^* = c^*(n)$ does not tend to a constant. The relatively few cases in which c^* is known to be a constant include 2-SAT, by Chvátal and Reed [3], Goerdts [12], and de la Vega [17], an extension to Max 2-SAT, by Coppersmith, Gamarnik, Hajiaghayi, and Sorkin [4], and the pure-literal threshold for a k -SAT formula, by Molloy [14].

Date: 10 June 2012.

[†] This research was supported by DIMACS, Center for Discrete Mathematics and Theoretical Computer Science, Rutgers, the State University of New Jersey, funded by the NSF under Grant No. DMS06-02942, Special Focus on Discrete Random Systems, and by the NSF under Grant No. DMS-0805996. The paper was prepared when the second author was a researcher in the Department of Mathematical Sciences, IBM T.J. Watson Research Center, Yorktown Heights NY 10598, USA. Journal submission April 2011, revised June 2012.

The most natural random model of the k -XORSAT problem is the “unconstrained” model in which each of the m equations’ k variables are drawn uniformly (without replacement) from the set of all n variables, and the right hand side values are uniformly 0 or 1; equivalently a random system $Ax = b$ is given by a matrix $A \in \{0, 1\}^{m \times n}$ drawn uniformly at random from the set of all such matrices with each row sum equal to k , and $b \in \{0, 1\}^m$ chosen uniformly at random.

The case $k = 2$ has been extensively studied. According to Kolchin [13] and Creignon and Daudé [5], for $c := 2m/n$ the random system has a solution with limiting probability $p(2m/n) + o(1)$, where $p(x) \in (0, 1)$ for $x < 1$, $p(1-) = 0$, and $p(x) \equiv 0$ for $x > 1$. Daudé and Ravelomanana [6], and Pittel and Yeum [15] analyzed the near-critical behavior of solvability probability for $2m/n = 1 + \varepsilon$, $\varepsilon = o(n^{-1/4})$.

For $k > 2$, Kolchin [13] analyzed the expected number of “critical row sets”, those with even column sums; their presence is necessary and sufficient for the (Boolean) rank of A to be less than m . He determined the thresholds c_k such that the expected number of critical sets goes to zero if $\lim m/n < c_k$ and to infinity if $\lim m/n > c_k$; in particular, $c_3 = 0.8894\dots$. Thus for $\lim m/n < c_k$, A is of full rank with probability tending to 1, implying $Ax = b$ is with high probability satisfiable, meaning that the solvability threshold c_k^* is at least c_k , $c_k^* \geq c_k$. On the other hand, the expected number of solutions of $Ax = b$ is 2^{n-m} , whence $\mathbb{P}(Ax = b \text{ is solvable})$ goes to zero for $\lim m/n > 1$, so $c_k^* \leq 1$. So Kolchin’s approach narrowed the location of the solvability threshold to the interval $[c_k, 1]$, but left open whether this threshold is actually c_k . (In fact, even the existence of a threshold remained unclear.) Even though the expected number of critical row sets goes to infinity for $\lim m/n > c_k$, it is still possible that the number of critical sets itself does not go to infinity in probability.

Dubois and Mandler [9] (see also [10]) introduced a “constrained” random k -XORSAT model, where b is still uniformly random, but A is uniformly random over the subset of matrices in which each column sum is at least 2. For $k = 3$ (3-XORSAT) they showed that its threshold is $\lim_{n,m \rightarrow \infty} m/n = 1$. They used this surprising result to determine the threshold $c_3^* \approx 0.9179 \in (c_3, 1)$ for the unconstrained model, thus completely answering Kolchin’s question for $k = 3$. Explicitly, they showed that

$$(1) \quad c_3^* = \frac{z^*}{3(1 - e^{-z^*})^2},$$

where z^* is the unique positive root of

$$(2) \quad (z - 3)e^z + 3 + 2z = 0.$$

The key idea for translating between the unconstrained and constrained models is that the 2-core of the 3-uniform hypergraph underlying a uniformly random unconstrained instance is a uniformly random constrained instance with a predictable density m/n , and the threshold value for the

unconstrained model is that for which the density of the core is 1. Dubois and Mandler derive a sharp asymptotic formula for the density of the core by analyzing a 2-core algorithm that successively eliminates any variable involved in no equations or just one, along with the associated equation if any. If the core is empty, then the initial system of equations $Ax = b$ is satisfiable. With high probability, $Ax = b$ remains solvable on the event “the core is nonempty, but its row density is below 1”. Dubois and Mandler concluded [9] with the suggestion that their methods could be extended to the general constrained k -XORSAT, $k \geq 3$. However, their approach — the second moment method for the number of solutions — requires solving a hard maximization problem with $\Theta(k)$ variables, a genuinely daunting task.

In this paper we prove that $\lim m/n = 1$ remains the solvability threshold for the general constrained k -XORSAT, for all $k \geq 3$. The crucial fact is that, for $k \geq 3$ and $\lim m/n < 1$, the expected number of critical row sets approaches 0 at a rate $n^{-(k-2)}$. What makes the asymptotic analysis doable is that in this, Kolchin-inspired, dual approach, the inevitable maximization problem is considerably simpler, involving just a *bivariate* function, no matter how large k is. We think that our argument provides a transparent explanation for why $\lim m/n = 1$ stubbornly remains the threshold for the constrained k -XORSAT for every $k \geq 3$.

Of course, once the constrained threshold of 1 is established for all $k \geq 3$, the reduction idea of Dubois and Mandler can be used to determine the unconstrained thresholds c_k^* . This time the issue is determination of the likely edge density for the 2-core of a random k -uniform hypergraph. Molloy computed this in [14], and we use his results to show that, for $k \geq 3$,

$$(3) \quad c_k^* = \frac{z_k^*}{k(1 - e^{-z_k^*})^{k-1}},$$

where z_k^* is the unique positive root of

$$(4) \quad (z - k)e^z + k + (k - 1)z = 0.$$

For $k = 3$, we are back to Dubois–Mandler’s (1)-(2).

While working on this project, we became aware (thanks to Alan Frieze) of a very recent extended abstract by Dietzfelbinger, Goerdt, Mitzenmacher, Montanari, Pagh, and Rink [7] on “cuckoo hashing”. It claims a proof of the constrained and unconstrained k -XORSAT thresholds, but the proof, influenced by [9], does not seem to be complete.

The remainder of the paper is organized as follows. In Section 2 we indicate that for the two-moments method for the number of solutions to be applicable in principle, it is necessary and sufficient that the expected number of non-empty critical row sets approaches zero. Thus we have a surprising reduction to the *first* order moment treatment of a dual characteristic of A , i. e. Kolchin’s approach. This is fortunate since the expected number of those row sets is quite amenable to asymptotic evaluation no matter how large k is.

In Section 3 first we show that, for the constrained model, instead of considering random 0–1 matrices A , it is asymptotically equivalent to consider uniformly random nonnegative integer matrices A subject to the same constraints on row sums ($= k$) and column sums (≥ 2). Second, using generating functions and Chernoff’s method, we obtain an exponential bound for the expected number of critical sets of any given cardinality. Third, we use this bound to show that, for $\lim m/n$ strictly below 1, and $k \geq 3$, the expected number of all critical sets is of order $n^{-(k-2)}$. Thus, there is unlikely to be any such set. Hence, with high probability, A is of full rank, and the system is solvable. On the other hand, since the expected number of solutions of $Ax = b$ is 2^{n-m} , for $\lim m/n > 1$ it tends to 0. So, with high probability, there are no solutions. We conclude that $\lim m/n = 1$ is a sharp threshold for solvability of $Ax = b$ in the constrained case for all $k \geq 3$.

Finally in Section 4 we give a derivation sketch for the unconstrained thresholds c_k^* , $k \geq 3$, based on Molloy’s results [14].

2. PROOF BACKGROUND

Let N be the number of solutions for a random formula, i.e., to the system of equations $Ax = b$ for a random pair (A, b) . By the first-moment method, the probability that a random formula is satisfiable satisfies

$$\mathbb{P}(Ax = b \text{ is satisfiable}) = \mathbb{P}(N > 0) \leq \mathbb{E}[N].$$

Let A be distributed arbitrarily over all $m \times n$ matrices, with 0,1-entries, and let b be independent of A , and uniformly distributed over $\{0,1\}^m$. As Dubois and Mandler [9] observed, regardless of the distribution of A ,

$$\mathbb{E}[N] = 2^{n-m}.$$

Since $\mathbb{E}[N] \rightarrow 0$ if $\lim m/n > 1$, this already establishes that, for any distribution of A , the satisfiability threshold c^* satisfies $c^* \leq 1$.

To show that $c^* \geq 1$ for $k = 3$ and A distributed uniformly over a constrained set of matrices, with all column sums 2 at least, [9] used the second-moment method based on

$$\mathbb{P}(Ax = b \text{ is satisfiable}) = \mathbb{P}(N > 0) \geq \mathbb{E}[N]^2 / \mathbb{E}[N^2].$$

The task was to show that $\mathbb{E}[N]^2 / \mathbb{E}[N^2] \rightarrow 1$ for $\lim m/n < 1$. Associating N^2 with the number of ordered pairs of solutions $\{x_1, x_2\}$ to $Ax = b$, they expressed $\mathbb{E}[N^2]$ as the sum of the expected numbers of those pairs x_1, x_2 over all possible values of the numbers of common components of x_1, x_2 and their locations in the equations. Determination of dominant contributors to that sum and demonstration that $\mathbb{E}[N^2]$ is indeed asymptotic to $\mathbb{E}[N]^2$ required an ingenious analysis of an attendant parametric maximization problem, with two variables and two parameters. Dubois and Mandler concluded with a remark that their method could be extended to k -XORSAT for any $k \geq 3$. However, our attempt to do that failed: the number of variables

grows linearly with k , making the corresponding parametric maximization exceedingly problematic.

To explain the approach we adopted in this paper, it is instructive first to derive an alternative representation for $\mathbb{E}[N^2]/\mathbb{E}[N]^2$, under the assumption that A and b are independent and b is uniform. By elementary linear algebra, for any fixed A , having rank $r(A)$ and nullity $n(A)$ over \mathcal{F}_2 , $Ax = b$ has $2^{m-r(A)}$ solutions for each of the $2^{r(A)}$ admissible values of b (those in $\{Ax : x \in \{0,1\}^n\}$), and 0 for the remaining values of b . Then, $\mathbb{E}[N^2] = \mathbb{E}[2^{2n-r(A)-m}]$, and

$$\mathbb{E}[N^2]/\mathbb{E}[N]^2 = \mathbb{E}[2^{m-r(A)}] = \mathbb{E}[2^{n(A^T)}].$$

Interpret $2^{n(A^T)}$, the number of solutions to $yA = 0$, as the number of *critical* row sets, that is, subsets $S \subseteq [m]$ such that all the column sums $\sum_{i \in S} a_{i,j}$, $j \in [n]$, are even; see Kolchin [13, Section 3.5]. Thus, defining

$$X = \# \text{ of non-empty critical row subsets of } A,$$

we have $\mathbb{E}[N^2]/\mathbb{E}[N]^2 = 1 + \mathbb{E}[X] \rightarrow 1$ iff $\mathbb{E}[X] \rightarrow 0$, *regardless of the distribution of A* . In words, the second moment method for N , the number of solutions, works iff $\mathbb{E}[X] \rightarrow 0$, where X is a dual parameter, the number of non-empty critical row sets of A . And it turns out that counting the expected number of critical sets is considerably easier than counting the expected number of pairs of solutions.

While we will not need it, it is worth mentioning that, under $\mathbb{E}[X] \rightarrow 0$, substantially more holds true. For if

$$\mathbb{E}[2^{m-r(A)} - 1] = \mathbb{E}[X] \rightarrow 0,$$

then $\mathbb{P}(r(A) = m) \rightarrow 1$, as $r(A) \leq m$. (This can happen only if $n \geq m$, of course.) In such a case, with high probability, for every b whence for the uniformly random b , we have $N = 2^{n-m}$, not just $N/2^{n-m} \rightarrow 1$ in probability.

We will show that indeed for $k \geq 3$ $\mathbb{E}[X] \rightarrow 0$, if $c := m/n$ satisfies $\lim_{m,n \rightarrow \infty} c < 1$, and A is uniformly distributed according to the constrained model. The inevitable calculus problem turns out to be relatively tractable. To be more specific, we show that the expected number of critical row sets of cardinality $\mu = \alpha n$ is bounded from above, roughly, by $\min_{\mathbf{z}} \exp[nH_k(\alpha, \mathbf{z}; c)]$, where H_k is defined in (18), and $\mathbf{z} = (z_1, z_2) > \mathbf{0}$. The task of finding the best $\mathbf{z} = \mathbf{z}(\alpha)$ and showing that $H_k(\alpha, \mathbf{z}(\alpha); c) < 0$ is hard, but we produce an explicit $\hat{\mathbf{z}}(\alpha)$ for which $\hat{h}_k(\alpha; c) := H_k(\alpha, \hat{\mathbf{z}}(\alpha); c)$ is proved to be negative.

3. COUNTING CRITICAL ROW SUBSETS

3.1. Probability spaces. Remember that throughout Section 3 we work with the constrained model. Let $\mathcal{A}_{m,n}$ denote the set of all $m \times n$ matrices with 0–1 entries, such that all m row sums are k , and all n column sums are

at least 2. For $\mathcal{A}_{m,n}$ to be nonempty, it is necessary that $km \geq 2n$. We will assume that $m, n \rightarrow \infty$ in such a way that $m = \Theta(n)$ and

$$(5) \quad \lim \left(k \frac{m}{n} - 2 \right) > 0.$$

The (Boolean) rank of a matrix $A \in \mathcal{A}_{m,n}$ is strictly less than m iff A has a non-empty critical row set.

A matrix $A \in \mathcal{A}_{m,n}$ may be interpreted as an outcome of the following allocation scheme. We have an $m \times n$ array of cells with k *indistinguishable* chips assigned to each of the m rows. For each row, the k chips are put in k distinct cells (so there is at most one chip per cell), subject to a constraint: every column gets at least two chips. Instead, we will consider the chips in every row to be *distinguishable*, so that each matrix A is obtained from $(k!)^m$ allocations (all that matters is whether a cell is occupied, not by which chip). Let $\mathcal{B}_{m,n}$ be the set of all feasible allocations in this scheme. Clearly, the uniform distribution on $\mathcal{A}_{m,n}$ is translated into that on $\mathcal{B}_{m,n}$. Let $\mathcal{C}_{m,n}$ be a relaxed version of $\mathcal{B}_{m,n}$: we do not require anymore that each of the mn cells gets at most one chip. Let B and C be distributed uniformly on $\mathcal{B}_{m,n}$ and $\mathcal{C}_{m,n}$, respectively. Crucially, and obviously, B is equal in distribution to C , conditioned on $C \in \mathcal{B}_{m,n}$.

To state a key lemma on $|\mathcal{A}_{m,n}|$, $|\mathcal{B}_{m,n}|$, and $|\mathcal{C}_{m,n}|$, we need some notation. Introduce

$$f(x) = \sum_{j \geq 2} \frac{x^j}{j!} = e^x - 1 - x.$$

By (5), for m, n large enough (which we assume from now on), $km/n > 2$ and is bounded away from 2. Any root of

$$(6) \quad \frac{xf'(x)}{f(x)} = \frac{km}{n} = kc$$

is a stationary point of $n \ln f(x) - km \ln x$. Since $\psi(x) := xf'(x)/f(x)$ is strictly increasing (see Note 1), and $\psi(0+) = 2$, $\psi(x) \sim x$ for $x \rightarrow \infty$, (6) has a unique root, which is bounded away from zero, the point where $n \ln f(x) - km \ln x$ attains its absolute maximum. Henceforth, let

$$\lambda = \lambda(kc)$$

be the unique root of (6).

Introduce a *truncated* Poisson random variable $Z = Z(\lambda)$,

$$\mathbb{P}(Z(\lambda) = j) = \frac{\lambda^j/j!}{f(\lambda)}, \quad j \geq 2.$$

Observe that the probability generating function (p.g.f.) of $Z(\lambda)$ is given by

$$\mathbb{E}[z^{Z(\lambda)}] = \frac{f(z\lambda)}{f(\lambda)};$$

in particular,

$$\mathbb{E}[Z(\lambda)] = \left. \frac{\partial}{\partial z} \frac{f(z\lambda)}{f(\lambda)} \right|_{z=1} = \frac{\lambda f'(\lambda)}{f(\lambda)} = \frac{km}{n},$$

see (6), and

$$\begin{aligned} (7) \quad \text{Var}[Z(\lambda)] &= \left. \frac{\partial^2}{\partial z^2} \frac{f(z\lambda)}{f(\lambda)} + \frac{\partial}{\partial z} \frac{f(z\lambda)}{f(\lambda)} - \left[\frac{\partial}{\partial z} \frac{f(z\lambda)}{f(\lambda)} \right]^2 \right|_{z=1} \\ &= \frac{\lambda^2 f''(\lambda)}{f(\lambda)} + \frac{\lambda f'(\lambda)}{f(\lambda)} - \left[\frac{\lambda f'(\lambda)}{f(\lambda)} \right]^2 = \Theta(\lambda), \end{aligned}$$

uniformly for $\lambda > 0$.

Note 1. That $\psi(x)$ strictly increases for $x \geq 0$ follows from

$$\psi'(x) = \frac{d}{dx} \frac{x f'(x)}{f(x)} = x^{-1} \text{Var}[Z(x)] > 0, \quad (x > 0).$$

Lemma 1.

$$(8) \quad |\mathcal{C}_{m,n}| = \frac{1 + O(n^{-1})}{\sqrt{2\pi n \text{Var}[Z(\lambda)]}} (km)! \frac{f(\lambda)^n}{\lambda^{km}},$$

$$(9) \quad \frac{|\mathcal{B}_{m,n}|}{|\mathcal{C}_{m,n}|} = \exp\left(-\frac{k-1}{2} \frac{\lambda e^\lambda}{e^\lambda - 1}\right) + o(1),$$

so that the fraction $|\mathcal{B}_{m,n}|/|\mathcal{C}_{m,n}|$ is bounded away from zero. Consequently

$$\begin{aligned} |\mathcal{A}_{m,n}| &= \frac{|\mathcal{B}_{m,n}|}{(k!)^m} = \frac{1 + o(1)}{\sqrt{2\pi n \text{Var}[Z(\lambda)]}} \frac{(km)!}{(k!)^m} \\ (10) \quad &\times \frac{f(\lambda)^n}{\lambda^{km}} \exp\left(-\frac{k-1}{2} \frac{\lambda e^\lambda}{e^\lambda - 1}\right). \end{aligned}$$

Note 2. For any $\mathcal{H} \subseteq \mathcal{B}_{m,n}$,

$$\begin{aligned} \mathbb{P}(B \in \mathcal{H}) &= \mathbb{P}(C \in \mathcal{H} \mid C \in \mathcal{B}_{m,n}) \\ &= \frac{\mathbb{P}(C \in \mathcal{H}, C \in \mathcal{B}_{m,n})}{|\mathcal{B}_{m,n}| / |\mathcal{C}_{n,m}|} \leq \frac{\mathbb{P}(C \in \mathcal{H})}{|\mathcal{B}_{m,n}| / |\mathcal{C}_{n,m}|}. \end{aligned}$$

Using this bound together with (9), we see that $\mathbb{P}(C \in \mathcal{H}) \rightarrow 0$ implies $\mathbb{P}(B \in \mathcal{H}) \rightarrow 0$. Replacing \mathcal{H} by its complement gives that $\mathbb{P}(C \in \mathcal{H}) \rightarrow 1$ implies $\mathbb{P}(B \in \mathcal{H}) \rightarrow 1$. In words, to show that an event in the sample space $\mathcal{B}_{m,n}$ is likely or unlikely, it suffices to prove that the event is likely or unlikely in the broader sample space $\mathcal{C}_{m,n}$, if both spaces are equipped with the uniform probability measure.

Proof of Lemma 1. Equation (10) is immediate from (8) and (9).

We first prove (8). To determine $|\mathcal{C}_{m,n}|$, recall that each row $i \in m$ is given its own k , mutually distinguishable, chips. This means all km chips are distinguishable, and each chip is affiliated with a row. Pool all km chips

together, and allocate them among n columns, subject to the constraint that every column gets at least two chips. Cell (i, j) gets those chips (if any) that are affiliated with row i and are allocated to column j .

We can get such an allocation by first permuting all the chips, then allocating the first $j_1 \geq 2$ chips to column 1, the next $j_2 \geq 2$ chips to column 2, *etc.*, noting that the permutation within each group is irrelevant. Adopting the notational convention that for $h(z) = \sum_j h_j z^j$, we define

$$[z^j] h(z) := h_j.$$

We thus have

$$\begin{aligned}
 |\mathcal{C}_{m,n}| &= \sum_{\substack{j_1 + \dots + j_n = km \\ j_1, \dots, j_n \geq 2}} \frac{(km)!}{j_1! \dots j_n!} \\
 &= (km)! [z^{km}] \left(\sum_{j \geq 2} \frac{z^j}{j!} \right)^n = (km)! [z^{km}] f(z)^n \\
 (11) \quad &= (km)! \frac{f(\lambda)^n}{\lambda^{km}} [z^{km}] \left(\frac{f(z\lambda)}{f(\lambda)} \right)^n = (km)! \frac{f(\lambda)^n}{\lambda^{km}} [z^{km}] (\mathbb{E}[z^{Z(\lambda)}])^n \\
 &= (km)! \frac{f(\lambda)^n}{\lambda^{km}} \mathbb{P} \left(\sum_{j=1}^n Z_j(\lambda) = km \right),
 \end{aligned}$$

where $Z_1(\lambda), \dots, Z_n(\lambda)$ are independent copies of $Z(\lambda)$. Since

$$\sum_{j=1}^n \mathbb{E}[Z_j(\lambda)] = n\mathbb{E}[Z(\lambda)] = km,$$

and $\lim \lambda > 0$, by a local limit theorem (Aronson et al [1], pages 174–176), the last probability is

$$\frac{1 + O(n^{-1})}{\sqrt{2\pi n \text{Var}[Z(\lambda)]}},$$

which proves (8).

We now prove (9). Let $C = \{c_{i,j}\}$ be distributed uniformly on $\mathcal{C}_{m,n}$. Let M denote the total number of cells that house 2 or more chips, i.e.

$$M = |\{(i, j) : c_{i,j} \geq 2\}|.$$

Let N be the total number of pairs of chips hosted by the same cell, i.e.,

$$N = \sum_{(i,j) : c_{i,j} \geq 2} \binom{c_{i,j}}{2} = \sum_{(i,j)} \binom{c_{i,j}}{2}.$$

$N = M$ iff there are no cells hosting more than 2 chips. Clearly

$$\frac{|\mathcal{B}_{m,n}|}{|\mathcal{C}_{m,n}|} = \mathbb{P}(C \in \mathcal{B}_{m,n}) = \mathbb{P}(N = 0).$$

Of course, $\mathbb{P}(N = 0) = \mathbb{P}(M = 0)$, but, unlike M , N is amenable to moment calculations.

Denoting the indicator of an event E by $\mathbf{1}(E)$, we write

$$N = \sum_{i \in [m], j \in [n]} \sum_{1 \leq u < v \leq k} \mathbf{1}(E(i, j; u, v)),$$

where $E(i, j; u, v)$ is the event that, out of k chips owned by row i , at least the two chips u and v , were put into the (i, j) cell. Each of these $mn \binom{k}{2}$ event indicators has the same expected value,

$$(12) \quad \mathbb{E}[E(i, j; u, v)] = (km - 2)! \frac{[x^{km-2}] f(x)^{n-1} e^x}{|\mathcal{C}_{m,n}|}.$$

To see this, note that once we have put two selected chips into an (i, j) -cell, we allocate the remaining $(km - 2)$ chips among n columns, at least two per column, except allowing the j th column to receive an unconstrained number of additional chips (it already has two), whence the sole e^x factor. Arguing as in (11),

$$(13) \quad [x^{km-2}] f(x)^{n-1} e^x = \frac{f(\lambda)^{n-1} e^\lambda}{\lambda^{km-2}} \mathbb{P} \left(\sum_{j=1}^{n-1} Z_j(\lambda) + \text{Po}(\lambda) = km - 2 \right),$$

where $\text{Po}(\lambda)$ stands for an independent, usual (not truncated) Poisson variable. The probability is again asymptotic to $(2\pi n \text{Var}[Z(\lambda)])^{-1/2}$. Therefore, using (8) and $km/n = \lambda f'(\lambda)/f(\lambda)$,

$$\mathbb{E}[N] = (1 + o(1)) \frac{mn \binom{k}{2}}{(km)_2} \frac{\lambda^2 e^\lambda}{f(\lambda)} = \gamma + o(1),$$

where

$$(14) \quad \gamma := \frac{k-1}{2} \frac{\lambda e^\lambda}{e^\lambda - 1}.$$

More generally, we claim that for every fixed $t \geq 1$ we have

$$(15) \quad \mathbb{E}[(N)_t] = \gamma^t + o(1),$$

where, as usual, $(a)_b := a(a-1) \cdots (a-b+1)$.

Let us prove (15). Denoting $\mathbf{i} = (i_1, \dots, i_t)$, $\mathbf{j} = (j_1, \dots, j_t)$, $\mathbf{u} = (u_1, \dots, u_t)$, $\mathbf{v} = (v_1, \dots, v_t)$, we have

$$(N)_t = \sum_{\mathbf{i} \in [m]^t, \mathbf{j} \in [n]^t} \sum_{\mathbf{u} < \mathbf{v}} \mathbf{1} \left(\bigcap_{s=1}^t E(i_s, j_s; u_s, v_s) \right).$$

Hence

$$\mathbb{E}[(N)_t] = \sum_{\mathbf{i} \in [m]^t, \mathbf{j} \in [n]^t} \sum_{\mathbf{u} < \mathbf{v}} \mathbb{P} \left(\bigcap_{s=1}^t E(i_s, j_s; u_s, v_s) \right).$$

We break the sum into two parts, Σ_1 and the remainder Σ_2 , where Σ_1 is the restriction to \mathbf{i} and \mathbf{j} , each having all its components distinct. In Σ_1 the total number of summands is $(m)_t(n)_t \binom{k}{2}^t$, and each summand is

$$(km - 2t)! \frac{[x^{km-2t}] f(x)^{n-t} (e^x)^t}{|\mathcal{C}_{m,n}|};$$

see the explanation following (12). Analogously to (13),

$$[x^{km-2t}] f(x)^{n-t} (e^x)^t = \frac{f(\lambda)^{n-t} (e^\lambda)^t}{\lambda^{km-2t}} \mathbb{P} \left(\sum_{j=1}^{n-t} Z_j(\lambda) + \sum_{s=1}^t \text{Po}_s(\lambda) = km - 2t \right),$$

where $\text{Po}_s(\lambda)$ are independent copies of $\text{Po}(\lambda)$, that are also independent of $Z_1(\lambda), \dots, Z_{n-t}(\lambda)$. As before, the probability is asymptotic to $(2\pi n \text{Var}[Z(\lambda)])^{-1/2}$. So, using (8) and recalling (14), we have

$$\begin{aligned} \Sigma_1 &\sim \frac{(m)_t(n)_t \binom{k}{2}^t}{(km)_{2t}} \left(\frac{\lambda^2 e^\lambda}{f(\lambda)} \right)^t \\ (16) \quad &\sim \left[\frac{mn \binom{k}{2}}{(km)^2} \frac{\lambda^2 e^\lambda}{f(\lambda)} \right]^t \rightarrow \gamma^t. \end{aligned}$$

In the case of Σ_2 , denoting $I = \{i_1, \dots, i_t\}$, $J = \{j_1, \dots, j_t\}$, we have $|I| + |J| \leq 2t - 1$. So the total number of attendant pairs (I, J) is at most $(m+n)^{2t-1} = O(m^{2t-1})$. And the number of pairs (\mathbf{i}, \mathbf{j}) inducing a given pair (I, J) is bounded above by a constant $s(t)$. For every one of those $s(t)$ choices, we select pairs of chips for each of the chosen t cells, in at most $\binom{k}{2}^t$ ways overall. Lastly, we allocate the remaining $(km - 2t)$ chips in such away that every column $j \in [n] \setminus J$ gets at least 2 chips. As in the case of Σ_1 , this can be done in

$$\begin{aligned} &(km - 2t)! [x^{km-2t}] f(x)^{n-|J|} (e^x)^{|J|} \\ &= \frac{f(\lambda)^{n-|J|} (e^\lambda)^{|J|}}{\lambda^{km-2t}} \mathbb{P} \left(\sum_{j=1}^{n-|J|} Z_j(\lambda) + \sum_{s=1}^{|J|} \text{Po}_s(\lambda) = km - 2t \right) \end{aligned}$$

number of ways. Again, the probability is asymptotic to $(2\pi n \text{Var}[Z(\lambda)])^{-1/2}$. So, as $e^\lambda > f(\lambda)$, the sum Σ_2 is of order

$$(17) \quad m^{2t-1} \frac{(km - 2t)!}{(km)!} \left(\frac{e^\lambda \lambda^2}{f(\lambda)} \right)^t = O(m^{2t-1}/m^{2t}) = O(m^{-1}).$$

Combining (16) and (17), and recalling (14), we conclude: for each fixed $t \geq 1$,

$$\mathbb{E}[(N)_t] = \gamma^t + o(1).$$

Therefore N is asymptotic, with all its moments and in distribution, to $\text{Po}(\gamma)$. In particular,

$$\mathbb{P}(N = 0) = \mathbb{P}(\text{Po}(\gamma) = 0) + o(1) = e^{-\gamma} + o(1).$$

This completes the proof of Lemma 1. \square

3.2. Main result. Armed with Lemma 1, we will prove our central result.

Theorem 2. *Suppose A is chosen uniformly at random (uar) from the set $\mathcal{A}_{m,n}$. Let $X_{m,n}$ denote the total number of non-empty critical row subsets of A . Suppose $k \geq 3$. If $m = m(n)$ is such that $\lim_{n \rightarrow \infty} c < 1$, then $\mathbb{P}(X_{m,n} = 0) \rightarrow 1$.*

Proof of Theorem 2. By Lemma 1, the discussion preceding it, and Note 2, we need only show that $\mathbb{E}[Y_{m,n}] \rightarrow 0$, where $Y_{m,n}$ is total number of non-empty critical row sets for the uniformly random matrix $C \in \mathcal{C}_{m,n}$. The proof rests on the following key claims.

Lemma 3. *Let $Y_{m,n}^{(\ell)}$ denote the total number of the critical row sets of cardinality ℓ , $2 \leq \ell \leq m$. Then, introducing $\mathbf{z} = (z_1, z_2) > \mathbf{0}$,*

$$(18) \quad \mathbb{E}[Y_{m,n}^{(\ell)}] \leq_b \sqrt{\frac{\lambda}{z_2}} \exp[nH_k(\alpha, \mathbf{z}; c)],$$

where

$$(19) \quad H_k(\alpha, \mathbf{z}; c) = -c(k-1)H(\alpha) + ck \ln \lambda - ck \ln z_1 - ck(1-\alpha) \ln z_2 + \ln \frac{f(z_1 + z_2) + f(z_2 - z_1)}{2f(\lambda)},$$

and

$$H(\alpha) := \alpha \ln \frac{1}{\alpha} + (1-\alpha) \ln \frac{1}{1-\alpha}, \quad (\text{entropy function}).$$

The reader certainly anticipates that, for a proper choice of \mathbf{z} , the exponential factor in (18) is dominant relative to $(\lambda/z_2)^{1/2}$.

Lemma 4. *Introduce $h_k(\alpha; c) := \inf_{\mathbf{z}} H_k(\alpha, \mathbf{z}; c)$. Then, for $k \geq 3$ and $c < 1$,*

$$(20) \quad h_k(\alpha; c) < 0, \quad \forall \alpha \in (0, 1].$$

More precisely, let $\alpha_k := \frac{e}{k^{\frac{k}{k-2}}}$; then

$$h_k(\alpha; c) \leq \begin{cases} c\alpha \ln \left[e \left(\frac{k}{e} \right)^{k/2} \alpha^{k/2-1} \right], & \alpha \in (0, 0.999\alpha_k], \\ -\beta, & \alpha \in [0.999\alpha_k, 1], \end{cases}$$

where $\beta = \beta(c, k) > 0$.

We will prove Lemma 4 by using $h_k(\alpha; c) \leq H_k(\alpha, \mathbf{z}; c)$ for a properly chosen $\mathbf{z} = \mathbf{z}(\alpha)$. For this choice of \mathbf{z} we will have $\lambda/z_2 = O(1)$ uniformly for α .

Proof of Lemma 3. By symmetry,

$$(21) \quad \mathbb{E}[Y_{m,n}^{(\ell)}] = \binom{m}{\ell} \mathbb{P}(\mathcal{D}_\ell); \quad \mathcal{D}_\ell := \bigcap_{j=1}^n \left\{ \sum_{i=1}^{\ell} c_{i,j} \text{ is even} \right\}.$$

By symmetry again,

$$(22) \quad \mathbb{P}(\mathcal{D}_\ell) = \sum_{\nu=1}^n \binom{n}{\nu} \mathbb{P}(\mathcal{D}_{\ell,\nu}),$$

$$\mathcal{D}_{\ell,\nu} := \bigcap_{j=1}^{\nu} \left\{ \sum_{i=1}^{\ell} c_{i,j} \text{ is even, positive} \right\} \cap \bigcap_{j=\nu+1}^n \left\{ \sum_{i=1}^{\ell} c_{i,j} = 0 \right\}.$$

Recalling that $\sum_{i \in [m]} c_{i,j} \geq 2$, we see that on the event $\mathcal{D}_{\ell,\nu}$

$$(23) \quad \sum_{i \leq \ell} c_{i,j} = \begin{cases} \text{even} > 0, & j \leq \nu, \\ 0, & j > \nu; \end{cases} \quad \sum_{i > \ell} c_{i,j} \geq \begin{cases} 0, & j \leq \nu, \\ 2, & j > \nu. \end{cases}$$

Thus on $\mathcal{D}_{\ell,\nu}$ the column sums of the two complementary submatrices, $\{c_{i,j}\}_{i \leq \ell, j \in [n]}$ and $\{c_{i,j}\}_{i > \ell, j \in [n]}$, are subject to the independent constraints.

Let $\mathcal{C}_{m,n}(\ell, \nu)$ denote the set of all matrices C with row sums k , which meet the constraints (23). Then $\mathbb{P}(\mathcal{D}_{\ell,\nu})$ is given by

$$(24) \quad p(\ell, \nu) := \mathbb{P}(\mathcal{D}_{\ell,\nu}) = \frac{|\mathcal{C}_{m,n}(\ell, \nu)|}{|\mathcal{C}_{m,n}|}.$$

By the independence of constraints on column sums for the upper and the lower submatrices of the matrices C in question,

$$|\mathcal{C}_{m,n}(\ell, \nu)| = a(\ell, \nu) \cdot b(m - \ell, \nu).$$

Here, just like $|\mathcal{C}_{m,n}|$, (1) $a(\ell, \nu)$ is the total number of ways to assign $k\ell$ chips among the first ν columns so that each of those columns gets a positive even number of chips, and (2) $b(m - \ell, \nu)$ is the total number of ways to assign $k(m - \ell)$ chips among all n columns so that each of the last $(n - \nu)$ columns gets at least 2 chips.

As in (11),

$$\begin{aligned} a(\ell, \nu) &= \sum_{\substack{j_1 + \dots + j_\nu = k\ell \\ j_s > 0, \text{ even}}} \frac{(k\ell)!}{j_1! \cdots j_\nu!} \\ &= (k\ell)! [z^{k\ell}] \left(\sum_{j > 0, \text{ even}} \frac{z^j}{j!} \right)^\nu = (k\ell)! [z^{k\ell}] (\cosh z - 1)^\nu, \end{aligned}$$

and

$$\begin{aligned} b(m - \ell, \nu) &= \sum_{\substack{j_1 + \dots + j_n = k(m - \ell) \\ j_1, \dots, j_\nu \geq 0; j_{\nu+1}, \dots, j_n \geq 2}} \frac{(k(m - \ell))!}{j_1! \dots j_n!} \\ &= (k(m - \ell))! [z^{k(m - \ell)}] (e^z)^\nu f(z)^{n - \nu}. \end{aligned}$$

Since the Taylor coefficients at $z = 0$ of both $(\cosh z - 1)^\nu$ and $e^{z\nu} f(z)^{n - \nu}$ are non-negative, we use these identities in a standard way to bound

$$\begin{aligned} (25) \quad a(\ell, \nu) &\leq (k\ell)! \frac{(\cosh z_1 - 1)^\nu}{z_1^{k\ell}}, \quad \forall z_1 > 0; \\ b(m - \ell, \nu) &\leq (k(m - \ell))! \frac{(e^{z_2})^\nu f(z_2)^{n - \nu}}{z_2^{k(m - \ell)}}, \quad \forall z_2 > 0. \end{aligned}$$

The guiding intuition is that, for z_1 and z_2 minimizing the RHS's of (25), the resulting bounds are the best possible, at least in terms of the ratio of corresponding logarithms. We also have

$$(26) \quad b(m - \ell, \nu) \leq_b (nz_2)^{-1/2} (k(m - \ell))! \frac{(e^{z_2})^\nu f(z_2)^{n - \nu}}{z_2^{k(m - \ell)}}, \quad \forall z_2 > 0.$$

(We use $A \leq_b B$ to indicate that $A = O(B)$, uniformly over all arguments in the expressions A and B . Later we will use this notation if, for every *fixed* k , $A = O(B)$, uniformly over all other arguments.) The bound (26) is stronger than the second bound in (25) if nz_2 is “very large”. (26) follows from the Cauchy integral formula

$$b(m - \ell, \nu) = \frac{(k(m - \ell))!}{2\pi} \oint_{\substack{z = z_2 e^{i\theta}: \\ \theta \in (-\pi, \pi]}} \frac{(e^z)^\nu f(z)^{n - \nu}}{z^{k(m - \ell)}} dz,$$

$|e^z| = e^{z_2} \exp[-z_2(1 - \cos \theta)]$, and less obvious

$$|f(z)| \leq |f(z_2)| \exp[-z_2(1 - \cos \theta)/3].$$

(See Pittel [16] for inequality, and [1] for how it works in combination with the Cauchy formula.)

Using (8), (24), first bound in (25), (26) and (8) from Lemma 1 together with (7), we obtain

$$p(\ell, \nu) \leq_b \sqrt{\frac{\lambda}{z_2}} \left(\frac{km}{k\ell} \right)^{-1} \frac{\lambda^{km}}{z_1^{k\ell} z_2^{k(m - \ell)}} \frac{[e^{z_2} (\cosh z_1 - 1)]^\nu f(z_2)^{n - \nu}}{f(\lambda)^n},$$

$\forall z_1 > 0, z_2 > 0$. Now, given n, m , the best z_1, z_2 certainly depend on both ℓ and ν . Lowering our sights, we are content to use z_1 and z_2 , whatever they will be, that depend on ℓ only. The reason is that combining this bound

and (22), (21), we obtain a bound

$$(27) \quad \mathbb{E}[Y_{m,n}^{(\ell)}] \leq_b \sqrt{\frac{\lambda}{z_2}} \binom{m}{\ell} \binom{km}{k\ell}^{-1} \lambda^{km} \times \frac{1}{z_1^{k\ell} z_2^{k(m-\ell)}} \left(\frac{f(z_2) + e^{z_2}(\cosh z_1 - 1)}{f(\lambda)} \right)^n, \quad \forall z_1, z_2 > 0,$$

that does not look forbiddingly complex. Observe that

$$f(z_2) + e^{z_2}(\cosh z_1 - 1) = \frac{f(z_1 + z_2) + f(z_2 - z_1)}{2}.$$

Introducing $\alpha = \ell/m$, $\mathbf{z} = (z_1, z_2)$, using the Stirling formula for factorials and exponentiating, we arrive at (18)-(19). \square

Proof of Lemma 4. First, as $f(x) > f(-x)$ for $x > 0$,

$$h_k(1; c) \leq H_k(1, (\lambda, 0); c) = \ln \frac{f(\lambda) + f(-\lambda)}{2f(\lambda)} < 0,$$

and it is bounded away from zero because $\lambda = \lambda(kc)$ is. So, by continuity of $h_k(\alpha; c)$ as a function of α , c ,

$$(28) \quad \sup_{\alpha \in [1-\delta, 1]} h_k(\alpha; c) < 0,$$

if $\delta = \delta(c, k) > 0$ is small enough. So we now focus on $\alpha \leq 1 - \delta_1$. Our task is to determine a proper $\hat{\mathbf{z}} = \hat{\mathbf{z}}(\alpha)$ for which $H_k(\alpha, \mathbf{z}; c) < 0$ for $\alpha \in (0, 1 - \delta_1]$ and $c < 1$. For guidance, let us assume that the infimum $h_k(\alpha; c)$ is attained at some $\mathbf{z}^*(\alpha; c) = (z_1^*(\alpha; c), z_2^*(\alpha; c))$. Such $\mathbf{z}^*(\alpha; c)$ is a solution of

$$(29) \quad \begin{aligned} \frac{\partial H_k(\alpha, \mathbf{z}; c)}{\partial z_1} &= -\frac{ck\alpha}{z_1} + \frac{f'(z_1 + z_2) - f'(z_2 - z_1)}{f(z_1 + z_2) + f(z_2 - z_1)} = 0, \\ \frac{\partial H_k(\alpha, \mathbf{z}; c)}{\partial z_2} &= -\frac{ck(1-\alpha)}{z_2} + \frac{f'(z_1 + z_2) + f'(z_2 - z_1)}{f(z_1 + z_2) + f(z_2 - z_1)} = 0. \end{aligned}$$

Consequently,

$$ck(1-2\alpha) = \frac{(z_2 - z_1)f'(z_1 + z_2) + (z_2 + z_1)f'(z_2 - z_1)}{f(z_1 + z_2) + f(z_2 - z_1)};$$

therefore $z_1^*(\alpha; c) < z_2^*(\alpha; c)$ for $\alpha < 1/2$, as the RHS is non-positive otherwise. For $\alpha \geq 1/2$, define

$$z_1^{**}(\alpha; c) = z_2^*(1 - \alpha; c), \quad z_2^{**}(\alpha; c) = z_1^*(1 - \alpha; c).$$

By (19), and $f(x) \leq f(|x|)$,

$$H_k(\alpha, \mathbf{z}^{**}(\alpha; c); c) \leq H_k(\alpha, \mathbf{z}^*(1 - \alpha; c); c);$$

hence it suffices to consider $\alpha \in (0, 1/2]$ only. Another easy corollary of (29) is

$$(30) \quad ck = \psi(z_1 + z_2) \cdot \frac{f(z_1 + z_2)}{f(z_1 + z_2) + f(z_2 - z_1)} + \psi(z_2 - z_1) \cdot \frac{f(z_2 - z_1)}{f(z_1 + z_2) + f(z_2 - z_1)},$$

$\psi(x) := xf'(x)/f(x)$. So $ck = \psi(\lambda)$ is a mean of the values of $\psi(x)$ for $x_1 = z_1 + z_2$ and $x_2 = z_2 - z_1$. Since $\psi(x)$ is increasing, we obtain

$$(31) \quad z_2^*(\alpha; c) - z_1^*(\alpha; c) < \lambda < z_1^*(\alpha; c) + z_2^*(\alpha; c).$$

It is also immediate from (29) and $z_2^*(\alpha; c) - z_1^*(\alpha; c) > 0$ that

$$(32) \quad \frac{\alpha}{z_1^*(\alpha; c)} \leq \frac{1 - \alpha}{z_2^*(\alpha; c)}.$$

It follows from (30) and (32) that $z_1^*(1/2; c) = z_2^*(1/2; c) = \lambda/2$. This observation suggests that, for $\alpha \leq 1/2$ and not too far from $1/2$, a good candidate for $\hat{\mathbf{z}}(\alpha)$ might be

$$(33) \quad \hat{z}_1(\alpha; c) = \lambda\alpha, \quad \hat{z}_2(\alpha; c) = \lambda(1 - \alpha).$$

An additional reason for this choice is that the resulting $H_k(\alpha, \hat{\mathbf{z}}(\alpha); c)$ given by (19) becomes

$$(34) \quad H_k(\alpha, \hat{\mathbf{z}}(\alpha; c); c) = cH(\alpha) + \ln \frac{f(\lambda) + f(\lambda(1 - 2\alpha))}{2f(\lambda)} \\ = \hat{H}_k(\alpha; \lambda) := k^{-1}\psi(\lambda)H(\alpha) + \ln \frac{f(\lambda) + f(\lambda(1 - 2\alpha))}{2f(\lambda)},$$

a relatively compact expression. Further, a close inspection of (29) shows that, for α small, $z_1^*(\alpha; c)$ and $z_2^*(\alpha; c)$ are asymptotic to $\lambda\Theta(\alpha^{1/2})$ and $\lambda(1 - \alpha)$ respectively. So let us *define*

$$\hat{z}_1(\alpha; c) = \lambda u \alpha^{1/2}, \quad \hat{z}_2(\alpha; c) = \lambda(1 - \alpha),$$

with $u = u(\alpha; c)$ to be determined. Notice that $\hat{z}_2 = \Theta(\lambda)$ for all $\alpha \leq 1/2$. By analogy with $\hat{\mathbf{z}}(\alpha; c)$, for $\alpha \in [1/2, 1 - \delta_1]$ we define

$$\hat{z}_1(\alpha; c) = \hat{z}_2(1 - \alpha; c), \quad \hat{z}_2(\alpha; c) = \hat{z}_1(1 - \alpha; c).$$

So, for all $\alpha \in (0, 1 - \delta_1]$, $\hat{z}_2(\alpha; c) = \Theta(\lambda)$, which makes the factor $(\lambda/\hat{z}_2)^{1/2}$ in (18) of order $O(1)$. And, like $\mathbf{z}^*(\alpha; c)$, it suffices to consider $\alpha \in (0, 1/2]$ only.

Let us use $H_k(\alpha, \hat{\mathbf{z}}(\alpha; c); c)$ as an upper bound for $h_k(\alpha; c)$. After cancellations, we have

$$(35) \quad H_k(\alpha, \hat{\mathbf{z}}(\alpha; c); c) = -c \left(\frac{k}{2} - 1 \right) \alpha \ln \frac{1}{\alpha} + c(1 - \alpha) \ln \frac{1}{1 - \alpha} - k c \alpha \ln u \\ + \ln \frac{f[\lambda(1 - \alpha + u \alpha^{1/2})] + f[\lambda(1 - \alpha - u \alpha^{1/2})]}{2f(\lambda)}.$$

Now, $\ln f(x)$ is convex, since

$$[\ln f(x)]'' = \frac{e^x(1-x-e^{-x})}{f^2(x)} < 0.$$

So, using $\psi(\lambda) = \lambda f'(\lambda)/f(\lambda) = kc$,

$$(36) \quad \frac{f[\lambda(1-\alpha \pm u\alpha^{1/2})]}{f(\lambda)} \leq \exp \left[\frac{\lambda f'(\lambda)}{f(\lambda)} (-\alpha \pm u\alpha^{1/2}) \right] \\ = e^{-kc\alpha} \cdot e^{\pm kcu\alpha^{1/2}}.$$

Combining this estimate with

$$\ln \cosh x \leq \frac{x^2}{2}, \quad (1-\alpha) \ln \frac{1}{1-\alpha} \leq \alpha,$$

we reduce (35) to a bound

$$H_k(\alpha, \hat{\mathbf{z}}(\alpha; c); c) \leq -c \left(\frac{k}{2} - 1 \right) \alpha \ln \frac{1}{\alpha} + c\alpha - kc\alpha \ln u \\ - kc\alpha - \frac{(kcu\alpha^{1/2})^2}{2}.$$

$u = u(\alpha; c)$, that minimizes this upper bound, is $(kc)^{-1/2}$, and we have

$$(37) \quad H_k(\alpha, \hat{\mathbf{z}}(\alpha; c); c) \leq c\alpha \ln \left[e \left(\frac{kc}{e} \right)^{k/2} \alpha^{k/2-1} \right].$$

Now, the RHS of (37) is negative for all $c \leq 1$ iff

$$(38) \quad \alpha \in (0, \alpha_k), \quad \alpha_k := \frac{e}{k^{\frac{k}{k-2}}};$$

($\alpha_k < 1/2$ for all $k \geq 3$.) Thus

$$(39) \quad h_k(\alpha; c) \leq H_k(\alpha, \mathbf{z}^*(\alpha; c); c) < 0, \quad \forall \alpha \in (0, \alpha_k).$$

For $\alpha \in [0.999\alpha_k, 1/2]$, we use $\hat{z}_1(\alpha; c)$, $\hat{z}_2(\alpha; c)$ defined in (33), with the corresponding $H_k(\alpha, \hat{\mathbf{z}}(\alpha; c); c)$ given by (34).

Lemma 5. (1) $\psi(x) = xf'(x)/f(x)$ is concave. (2) Consequently, if $\alpha \in (0, 1/2]$, $\lambda > 0$ are such that $\hat{H}_k(\alpha; \lambda) \geq 0$, then $\partial \hat{H}_k(\alpha; \lambda)/\partial \lambda > 0$. In words, as a function of λ , $\hat{H}_k(\alpha; \lambda)$ is strictly increasing when it is non-negative.

We will prove Lemma 5 at the end. The part (2) of this Lemma implies an important corollary.

Corollary 6. For $c < 1$,

$$(40) \quad \{ \alpha \in (0, 1/2] : H_k(\alpha, \hat{\mathbf{z}}(\alpha; c); c) \geq 0 \} \\ \subset \{ \alpha \in (0, 1/2] : H_k(\alpha, \hat{\mathbf{z}}(\alpha; 1); 1) > 0 \}.$$

Equivalently, $H_k(\alpha, \hat{\mathbf{z}}(\alpha; c); c)$ is negative for all α 's such that $\hat{H}_k(\alpha; \lambda_k) < 0$.

Proof of Corollary 6. If for some $c < 1$ and $\alpha \in (0, 1/2]$, we have $H_k(\alpha, \hat{\mathbf{z}}(\alpha; c); c) \geq 0$, then $\hat{H}_k(\alpha; \lambda(kc)) \geq 0$, where $\psi(\lambda(kc)) = kc$. By Lemma 5, $\hat{H}_k(\alpha; \lambda) > 0$ for all $\lambda > \lambda(kc)$, and, in particular, for $\lambda_k = \lambda(kc)|_{c=1}$, which means that $H_k(\alpha, \hat{\mathbf{z}}(\alpha; 1); 1) > 0$. \square

As $f'(x)/f(x) > 1$, it follows that $\lambda_k < k$. By induction on k , $\lambda_k > k - 1$ for all $k \geq 3$. In addition, $\psi(\lambda_k) = k$ is equivalent to

$$\lambda_k + \frac{\lambda_k^2}{f(\lambda_k)} = k,$$

so that

$$(41) \quad \frac{\lambda_k^2}{f(\lambda_k)} < 1,$$

which will come handy in a moment. Since

$$H''(\alpha) = -\frac{1}{\alpha(1-\alpha)} \leq -4,$$

we have

$$H(\alpha) - H(1/2) = H(\alpha) - \ln 2 \leq -\frac{1}{2}(1-2\alpha)^2.$$

Also, as $f(x) \leq e^x x^2/2$ for $x \geq 0$,

$$f(\lambda(1-2\alpha)) \leq e^{\lambda(1-2\alpha)} \frac{\lambda^2(1-2\alpha)^2}{2}, \quad \alpha \leq 1/2.$$

So

$$\begin{aligned} \hat{H}_k(\alpha; \lambda_k) &= H(\alpha) - \ln 2 + \ln \left[1 + \frac{f(\lambda_k(1-2\alpha))}{f(\lambda_k)} \right] \\ &\leq -\frac{1}{2}(1-2\alpha)^2 \left[1 - \frac{\lambda_k^2}{f(\lambda_k)} e^{\lambda_k(1-2\alpha)} \right] < 0, \end{aligned}$$

if

$$(42) \quad \alpha > \tilde{\alpha}_k := \frac{1}{2} \left[1 - \frac{1}{\lambda_k} \ln \frac{f(\lambda_k)}{\lambda_k^2} \right];$$

$\tilde{\alpha}_k < 1/2$ by (41), and $\tilde{\alpha}_k > 0$ as $f(\lambda_k)/\lambda_k^2 < e^{\lambda_k}$. By log-convexity of $f(x)$, and $\psi(\lambda_k) = k$,

$$\frac{f(\lambda_k(1-2\alpha))}{f(\lambda_k)} < \exp \left[\frac{f'(\lambda_k)}{f(\lambda_k)} (-\lambda_k 2\alpha) \right] = e^{-2\alpha k}.$$

So, for $\alpha \leq \tilde{\alpha}_k$ we have

$$(43) \quad \hat{H}_k(\alpha; \lambda_k) < H(\tilde{\alpha}_k) + \ln \frac{1 + e^{-2\alpha k}}{2}.$$

To simplify this bound, let us show that $\tilde{\alpha}_k$ decreases with k . We write

$$\left(\frac{1}{x} \ln \frac{f(x)}{x^2} \right)' = -\frac{1}{x^2} G(x),$$

where

$$G(x) = \ln \frac{f(x)}{x^2} + 2 - \frac{xf'(x)}{f(x)}.$$

Now $G(+0) = -\ln 2 < 0$, and

$$G'(x) = \frac{1}{x} \left[\frac{xf'(x)}{f(x)} - 2 - x \left(\frac{xf'(x)}{f(x)} \right)' \right] < 0,$$

since

$$\lim_{x \rightarrow 0} \frac{xf'(x)}{f(x)} = 2,$$

and $xf'(x)/f(x)$ is concave. Therefore $G(x) < 0$ for $x > 0$, so that $x^{-1} \ln(f(x)/x^2)$ increases with x . So, by (42), $\tilde{\alpha}_k$ indeed decreases with k . So, considering $k \geq 7$, we replace (43) by a cruder bound

$$(44) \quad \hat{H}_k(\alpha; \lambda_k) < H(\tilde{\alpha}_7) + \ln \frac{1 + e^{-2\alpha k}}{2}, \quad \forall \alpha \leq \tilde{\alpha}_7.$$

Numerically,

$$\lambda_7 \approx 6.953, \quad \tilde{\alpha}_7 \approx 0.2794, \quad H(\tilde{\alpha}_7) \approx 0.5924.$$

By (44), $\hat{H}_k(\alpha; \lambda_k) < 0$ if

$$\alpha \in \left(\frac{\beta_7}{k}, \frac{1}{2} \right), \quad \beta_7 := \frac{1}{2} \ln \left(\frac{1}{2e^{-H(\tilde{\alpha}_7)} - 1} \right) \approx 1.122.$$

If $0.999\alpha_k$, (see (38) for α_k), exceeds β_7/k for $k \geq 7$, then—in the light of (39)—we will be able conclude that

$$(45) \quad H_k(\alpha, \mathbf{z}^*(\alpha; c); c) < 0, \quad \forall \alpha \in (0, 1/2], \quad c < 1.$$

And, in fact, the inequality

$$\frac{\beta_7}{k} < 0.999\alpha_k = 0.999 \frac{e}{k^{\frac{k}{k-2}}},$$

does hold for $k \geq 7$, since

$$\inf \left\{ \frac{ke}{k^{\frac{k}{k-2}}} : k \geq 7 \right\} = \frac{7e}{7^{7/5}} = \frac{e}{7^{2/5}} \approx 1.248 > 1.123 \approx \frac{\beta_7}{0.999}.$$

(Needless to say, 7 happens to be the first integer $j \geq 3$ for which

$$\frac{\beta_j}{k} < 0.999\alpha_k, \quad \forall k \geq j.)$$

The values $k = 6, 5, 4$ can be dealt with by pushing the above device a bit further. See Appendix (a). Thus, contingent on Lemma 5, we have proved the $k > 3$ part of Lemma 4.

Proof of Lemma 5. (1) It suffices to prove convexity of $\phi(x) = x^2/f(x)$, since

$$\psi(x) := \frac{xf'(x)}{f(x)} = x + \phi(x).$$

Now

$$\begin{aligned}\phi''(x) &= \left(\frac{2x}{f(x)} - \frac{x^2 f'(x)}{f(x)^2} \right)' \\ &= \frac{2}{f(x)} - \frac{4x f'(x)}{f(x)^2} - \frac{x^2 f''(x)}{f(x)^2} + \frac{2x^2 (f'(x))^2}{f(x)^3} \\ &= \frac{g(x)}{f(x)^3}.\end{aligned}$$

Here

$$\begin{aligned}g(x) &= 2f(x)^2 - 4x f'(x) f(x) - x^2 f''(x) f(x) + 2x^2 (f'(x))^2 \\ &= \sum_{j \geq 6} g_j x^j,\end{aligned}$$

and

$$g_j = \frac{j-1}{j!} [2^{j-2}(j-8) + j^2 - j + 4] > 0, \quad \forall j \geq 6.$$

Therefore $\phi''(x) > 0$ for all $x > 0$.

(2) By the definition of $\hat{H}_k(\alpha; \lambda)$ in (34),

$$\begin{aligned}\frac{\partial \hat{H}_k(\alpha; \lambda)}{\partial \lambda} &= \frac{1}{k} \left(\frac{\lambda f'(\lambda)}{f(\lambda)} \right)' H(\alpha) + \frac{\left(\frac{f(\lambda(1-2\alpha))}{f(\lambda)} \right)'_{\lambda}}{1 + \frac{f(\lambda(1-2\alpha))}{f(\lambda)}} \\ &= \frac{1}{k} \psi'(\lambda) H(\alpha) + \lambda^{-1} \frac{\frac{f(\lambda(1-2\alpha))}{f(\lambda)}}{1 + \frac{f(\lambda(1-2\alpha))}{f(\lambda)}} \cdot [\psi(\lambda(1-2\alpha)) - \psi(\lambda)].\end{aligned}$$

Here, by concavity of $\psi(\cdot)$,

$$\psi(\lambda(1-2\alpha)) - \psi(\lambda) \geq -2\alpha \lambda \psi'(\lambda),$$

and, by the condition $\hat{H}_k(\alpha; \lambda) \geq 0$,

$$H(\alpha) \geq \frac{k}{\psi(\lambda)} \ln \frac{2f(\lambda)}{f(\lambda) + f(\lambda(1-2\alpha))} = \frac{k}{\psi(\lambda)} \ln \frac{2}{1 + \frac{f(\lambda(1-2\alpha))}{f(\lambda)}}.$$

So, since $\psi'(\lambda) > 0$, we have $\partial \hat{H}_k(\alpha; \lambda) / \partial \lambda > 0$, if

$$(46) \quad \ln \frac{2}{1 + \frac{f(\lambda(1-2\alpha))}{f(\lambda)}} - 2\alpha \psi(\lambda) \frac{\frac{f(\lambda(1-2\alpha))}{f(\lambda)}}{1 + \frac{f(\lambda(1-2\alpha))}{f(\lambda)}} > 0.$$

Recalling log-convexity of $f(\cdot)$, we also have

$$\frac{f(\lambda(1-2\alpha))}{f(\lambda)} \leq \exp[-2\alpha \psi(\lambda)].$$

So (46) definitely holds if

$$\ln \frac{2}{1 + e^{-z}} - z \frac{e^{-z}}{1 + e^{-z}} > 0, \quad \forall z > 0.$$

And indeed, using convexity of $\ln(\cdot)$, we get

$$\ln \frac{2}{1+e^{-z}} > \frac{1-e^{-z}}{2} > \frac{ze^{-z}}{1+e^{-z}},$$

the last inequality being equivalent to $\sinh z > z$. This concludes the proof of Lemma 5, whence of Lemma 4. \square

The next claim follows immediately from Lemma 19 where $z_2 = \Theta(\lambda)$, and Lemma 4.

Corollary 7. *Let $\varepsilon = 0.999\alpha_k$. For every $k \geq 3$ and $c \leq 1$, there exists $\gamma_0 = \gamma_0(c, k) > 0$ such that, for $2 \leq \ell \leq \varepsilon m$,*

$$\mathbb{E}[Y_{m,n}^{(\ell)}] \leq_b \left(\frac{m\gamma_0}{\ell} \right)^{-\ell(k/2-1)}.$$

Consequently,

$$\sum_{\ell=2}^{\lfloor \varepsilon m \rfloor} \mathbb{E}[Y_{m,n}^{(\ell)}] = O(m^{-(k-2)}).$$

Furthermore, for $c < 1$

$$\sum_{\ell \geq \varepsilon m} \mathbb{E}[Y_{m,n}^{(\ell)}] = O(e^{-n\beta}),$$

$$\beta = \beta(c, k) > 0.$$

We conclude that, for $c < 1$,

$$\sum_{\ell=2}^m \mathbb{E}[Y_{m,n}^{(\ell)}] = O(m^{-(k-2)}).$$

The $k > 3$ part of Theorem 2 is proved. \square

The steady rise of technicality from $k \geq 7$ to $k = 6, 5, 4$, in that order, makes it clear that the case $k = 3$, dealt with differently in [9], would likely be more challenging for our approach. Here is a modification of our argument above necessary to cover $k = 3$ as well. This time we are more dependent on computer than for $k \geq 4$. Like $k \geq 4$, we need to show that, for $c < 1$, $h_3(\alpha; c) < 0$ for $\alpha \in (0, 1/2]$. Beginning with arbitrary $k \geq 3$, we define

$$(47) \quad \begin{aligned} \bar{z}_1 &= \bar{z}\alpha^{\frac{k-1}{k}} & \bar{z}_2 &= \bar{z}(1-\alpha)^{\frac{k-1}{k}}, \\ \bar{z} &= \lambda \exp[-k^{-1}H(\alpha)]; \end{aligned}$$

in particular, $z_2 = \Theta(\lambda)$. Notice that letting $k \uparrow \infty$ we get the equations (33). What makes this selection of controllable z_1, z_2 promising is an observation that for the exponent defined in (19) we have

$$(48) \quad \begin{aligned} H_k(\alpha, \bar{z}; c) &= cH(\alpha) + \ln \frac{f(\bar{z}_1 + \bar{z}_2) + f(\bar{z}_2 - \bar{z}_1)}{2f(\lambda)} \\ &= \bar{H}_k(\alpha; \lambda) := \frac{\lambda f'(\lambda)}{kf(\lambda)} H(\alpha) + \ln \frac{f(\bar{z}_1 + \bar{z}_2) + f(\bar{z}_2 - \bar{z}_1)}{2f(\lambda)}, \end{aligned}$$

a formula nearly as compact as (34) for $z_1 = \alpha\lambda$, $z_2 = (1 - \alpha)\lambda$. (Less formally, the optimal $(z_1^*(\alpha; c), z_2^*(\alpha; c))$ happens to be given by (47) at *stationary points* of $h_k(\alpha; c) = \inf_{\mathbf{z}} H_k(\alpha, \mathbf{z}; c)$, which is considered as a function of α and of a *continuously* varying parameter k .) Moreover, in a crucial analogy to $\hat{H}_k(\alpha; \lambda)$,

$$(49) \quad \frac{\partial}{\partial \lambda} \bar{H}_k(\alpha; \lambda) \geq 0, \quad \text{if } \bar{H}_k(\alpha; \lambda) \geq 0.$$

The proof of (49) runs parallel to that of Lemma 5 but more involved, see Appendix (b).

Consequently, for $k = 3$, we need only show that $\bar{H}_3(\alpha; \lambda_3) < 0$ for all $\alpha \in (0, 1/2)$. Now, it follows from (47) and (48) that

$$\begin{aligned} \bar{H}_3(\alpha; \lambda_3) &= -2\alpha + O(\alpha^{4/3}), \quad \alpha \downarrow 0; \\ \bar{H}_3(\alpha; \lambda_3) &= -\gamma(1/2 - \alpha)^2 + O((1/2 - \alpha)^3), \quad \alpha \uparrow 1/2; \\ \gamma &:= \frac{4}{3} - \frac{8}{9} \frac{\lambda_3^2}{f(\lambda_3)} \approx 0.577 > 0. \end{aligned}$$

Thus $\bar{H}_3(0; \lambda_3) = \bar{H}_3(1/2; \lambda_3) = 0$, and $\bar{H}_3(\alpha; \lambda_3) < 0$ for $\alpha \neq 0, 1/2$ either close to 0 or to $1/2$. Using Maple we have checked that the *univariate* function $\bar{H}_3(\alpha; \lambda_3)$ is negative for all $\alpha \in (0, 1/2)$.

The proof of Theorem 2 is complete. \square

4. SOLVABILITY THRESHOLD FOR THE UNCONSTRAINED k -XORSAT

As we mentioned in Introduction, Dubois and Mandler had determined the threshold for the unconstrained 3-XORSAT by analyzing the terminal state of a reduction process applied to an instance of this problem. At each step of this process a variable appearing in only one equation is deleted together with the row corresponding to that equation. In graph-theoretical terms, this is a process that delivers a 2-core of the underlying 3-uniform hypergraph. So the solvability threshold c_3^* is the edge density $\lim m/n$ for the unconstrained instance for which the edge density of its 2-core is 1. Dietzfelbinger, Goerdts, Mitzenmacher, Montanari, Pagh, and Rink [7, 8] demonstrated that the known results on 2-cores of the random k -uniform hypergraph could be used for analogous treatment of unconstrained k -XORSAT in general, for $k \geq 3$, if the constrained threshold was known to be 1. Let us show how the rigorous analysis of the cores by Molloy [14] delivers a short derivation of c_k^* , for all $k \geq 3$.

Let

$$\hat{c} = k^{-1} \min_{x>0} \frac{x}{(1 - e^{-x})^{k-1}}, \quad k \geq 3.$$

From [14, Theorem 1] it follows that the terminal matrix obtained via the reduction process from the random $m \times n$ matrix A , i. e. the 2-core of the underlying hypergraph, is with high probability (whp) empty if $\lim m/n < \hat{c}$. Hence, for $\lim m/n < \hat{c}$, whp $Ax = b$ is solvable.

If on the other hand $c := \lim m/n > \hat{c}$, then whp the terminal matrix has $\mathcal{N} = (1 - e^x)n + o(n)$ columns, where $x = x(c)$ is the largest solution to

$$(50) \quad c = k^{-1} \frac{x}{(1 - e^{-x})^{k-1}}.$$

Importantly, $x(c)$ strictly increases with c . In this case a variable (column) chosen uniformly at random from among the \mathcal{N} columns of the terminal matrix appears in a random number of equations (rows) of the terminal matrix whose asymptotic distribution is a Poisson $Z(x)$ conditioned on the event $\{Z(x) \geq 2\}$. (See shortly before (7) for a reminder.) Therefore, denoting by \mathcal{M} the number of rows of the terminal matrix, we obtain

$$k \mathbb{E}[\mathcal{M}] \sim \mathbb{E}[\mathcal{N}] \cdot \mathbb{E}[Z(x) \mid Z(x) \geq 2].$$

Since \mathcal{M} and \mathcal{N} are sharply concentrated around $\mathbb{E}[\mathcal{M}]$ and $\mathbb{E}[\mathcal{N}]$ respectively, we obtain that, in probability, the terminal edge density is asymptotic to

$$k^{-1} \mathbb{E}[Z(x) \mid Z(x) \geq 2] = k^{-1} \frac{x \mathbb{P}(Z(x) \geq 1)}{\mathbb{P}(Z(x) \geq 2)} = k^{-1} \frac{x(e^x - 1)}{e^x - 1 - x}.$$

We emphasize that, conditioned on the row column set having cardinalities \mathcal{M}, \mathcal{N} respectively, the terminal matrix is distributed uniformly on the set $\mathcal{A}_{\mathcal{M}, \mathcal{N}}$, if we label rows and columns by the elements of $[\mathcal{M}]$ and $[\mathcal{N}]$. Let x_k be the unique positive root of

$$k^{-1} \frac{x_k(e^{x_k} - 1)}{e^{x_k} - 1 - x_k} = 1 \iff (x_k - k)e^{x_k} + k + (k - 1)x_k = 0.$$

It follows then from (50) that c_k^* , the threshold edge density of the unconstrained matrix A is given by

$$c_k^* = k^{-1} \frac{x_k}{(1 - e^{-x_k})^{k-1}}.$$

The equations (3)-(4) are proved. \square

Acknowledgment. We are grateful to Alan Frieze for sustaining our interest in this project. Our sincere thanks go to the anonymous referees for their effort and time and for many penetrating critical comments that gave us an invaluable opportunity to improve the paper. One of the referees, in addition to a series of insightful suggestions, pinpointed a subtle oversight in the proof of Lemma 1, for which we are genuinely grateful. Our proof that $h_k(\alpha; c) < 0$ had been limited to boundary values of the parameters, with numerical experiments non-rigorously establishing negativity in the interior. The editors passed to us a proof sketch by Paul Balister, using patch-work approximation in various parameter regimes to get finitely away from the boundaries, coupled with a proposal to use interval arithmetic for the interior. We adopted the patch-work approach, differently implemented, to get a proof which is computation-free for $k \geq 7$, and has minimal computation (and only for univariate functions) for smaller k . We are indebted to Paul for his inspiring help.

REFERENCES

- [1] J. Aronson, A. Frieze and B. Pittel, On maximum matching in sparse random graphs: Karp-Sipser revisited, *Random Struct. Algorithms*, **12** (1998), 111–177.
- [2] B. Bollobás, C. Borgs, J. T. Chayes, Jeong-Han Kim, and D. B. Wilson, The scaling window of the 2-SAT transition, *Random Struct. Algorithms*, **18** (2001) 201–256.
- [3] V. Chvátal and B. Reed, Mick gets some (the odds are on his side), *33th Annual Symposium on Foundations of Computer Science (Pittsburgh, PA, 1992)*, *IEEE Comput. Soc. Press, Los Alamitos, CA* (1992) 620–627.
- [4] D. Coppersmith, D. Gamarnik, M. T. Hajiaghayi and G. B. Sorkin, Random MAX SAT, random MAX CUT, and their phase transitions, *Random Struct. Algorithms* **24** (2004) 502–545.
- [5] N. Creignon and H. Daudé, Smooth and sharp thresholds for random k -XOR-CNF satisfiability, *Theor. Inform. Appl.* **37** (2003) 127–147.
- [6] H. Daudé and V. Ravelomanana, Random 2-XORSAT at the satisfiability threshold, *LATIN 2008: Theoretical Informatics, 8th Latin American Symposium Proceedings* (2008) 12–23.
- [7] M. Dietzfelbinger, A. Goerdt, M. Mitzenmacher, A. Montanari, R. Pagh, and M. Rink, Tight thresholds for cuckoo hashing via XORSAT, *Proceedings of the 37th international colloquium conference on Automata, languages and programming (ICALP'10)*, S. Abramsky, C. Gavoille, C. Kirchner, F. M. Auf Der Heide, and P. G. Spirakis (Eds.) (Springer-Verlag, Berlin, Heidelberg) (2010) 213–225.
- [8] M. Dietzfelbinger, A. Goerdt, M. Mitzenmacher, A. Montanari, R. Pagh and M. Rink, Tight thresholds for cuckoo hashing via XORSAT, *arXiv:0912.0287v3* (2010).
- [9] O. Dubois and J. Mandler, The 3-XORSAT threshold, *Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science, FOCS 2002 (Vancouver, BC, Canada)*, *IEEE Computer Society* (2002) 769–778.
- [10] O. Dubois and J. Mandler, The 3-XORSAT threshold, *C. R. Acad. Sci. Paris, Ser. I* **335** (2002) 963–966.
- [11] E. Friedgut, Necessary and sufficient conditions for sharp thresholds of graph properties, and the k -SAT problem, *J. Amer. Math. Soc.* **12** (1999), 1017–1054.
- [12] A. Goerdt, A threshold for unsatisfiability, *J. Comput. System Sci.* **53** (1996) 469–486.
- [13] V. F. Kolchin, Random graphs, *Encyclopedia of Mathematics and its Applications*, vol. 53, *Cambridge University Press, Cambridge* (1999).
- [14] M. Molloy, Cores in random hypergraphs and Boolean formulas, *Random Struct. Algorithms* **27** (2005) 124–135.
- [15] B. Pittel and Ji-A Yeum, How frequently is a system of 2-linear equations solvable? *Electronic J. Combin.* **17** (2010) # R 92.
- [16] B. Pittel, Paths in a Random Digital Tree: Limiting Distributions, *Adv. Appl. Prob.* **18** (1986) 139–155.
- [17] W. Fernandez de la Vega, On random 2-SAT, manuscript (1992).

Appendix (a). $k = 6, 5, 4$ cases.

Let $k = 6$. Define $x_0 = \tilde{\alpha}_6$, and recursively

$$(51) \quad x_j = \frac{1}{2 \cdot 6} \ln \left(\frac{1}{2e^{-H(x_{j-1})} - 1} \right), \quad j \geq 1.$$

The recurrence (51) is equivalent to

$$H(x_{j-1}) + \ln \frac{1 + e^{-2 \cdot 6 x_j}}{2} = 0.$$

$H(x)$ is strictly increasing on $[0, 1/2]$, and $\ln(1+e^{-12x})$ is strictly decreasing. So if $x_j < x_{j-1}$ for some $j \geq 1$ then

$$H(x) + \ln \frac{1 + e^{-2 \cdot 6x}}{2} < 0, \quad \forall x \in [x_j, x_{j-1}].$$

Consequently, if $x_0 > x_1 > \dots > x_j$ then

$$(52) \quad \hat{H}_6(\alpha; \lambda_6) < 0, \quad \forall \alpha \in [x_j, x_0] = [x_j, \tilde{\alpha}_6],$$

whence

$$\hat{H}_6(\alpha; \lambda_6) < 0, \quad \forall \alpha \in [x_j, 1/2].$$

Numerically,

$$x_0 = \tilde{\alpha}_6 \approx 0.302, \quad x_1 \approx 0.207, \quad x_2 \approx 0.134,$$

i. e. $x_0 > x_1 > x_2$ indeed. And we see that, by the definition of α_6 in (38),

$$x_2 < 0.180 < 0.185 \approx 0.999\alpha_6 = 0.999 \frac{e}{6^{\frac{6}{6-2}}}.$$

Combining (52) and (39) we obtain that

$$H_6(\alpha, \hat{\mathbf{z}}(\alpha; c); c) < 0, \quad \forall \alpha \leq 1/2, \quad c < 1.$$

Let $k = 5$. Define $\{x_j\}$ by the recurrence obtained from (51), with 6 replaced by 5, and $x_0 = \tilde{\alpha}_5$. This time we get $x_0 > x_1 > x_2 > x_3 \approx 0.177$, while $\alpha_5 \approx 0.186$. So, as $x_3 < 0.999\alpha_5$, we have

$$H_5(\alpha, \hat{\mathbf{z}}(\alpha; c); c) < 0, \quad \forall \alpha \leq 1/2, \quad c < 1.$$

Let $k = 4$. Here

$$x_0 = \tilde{\alpha}_4 \approx 0.374 > \alpha_4 = \frac{e}{16},$$

but $x_1 \approx 0.429 > x_0$. In this case we introduce a sequence $\{y_j\}$, such that $y_0 = 0.999\alpha_4 \approx 0.170$ and, for $j \geq 1$,

$$H(y_j) + \ln \frac{f(\lambda_4) + f(\lambda_4(1 - 2y_{j-1}))}{2f(\lambda_4)} = 0,$$

(cf. the definition of $\hat{H}_k(\alpha; \lambda_k)$ in (34)), or equivalently

$$(53) \quad y_j = H^{-1} \left[\ln \frac{2f(\lambda_4)}{f(\lambda_4) + f(\lambda_4(1 - 2y_{j-1}))} \right].$$

Here $H^{-1}(x)$ is inverse to the entropy function, defined for $x \in [0, \ln 2]$. The function

$$\ln \frac{f(\lambda_4) + f(\lambda_4(1 - 2y))}{2f(\lambda_4)}$$

is strictly decreasing. Therefore, as before, if $y_0 < y_1 < \dots < y_j$ for some $j > 0$, then

$$\hat{H}_k(\alpha; \lambda_k) = H(\alpha) + \ln \frac{f(\lambda_4) + f(\lambda_4(1 - 2\alpha))}{2f(\lambda_4)} < 0, \quad \forall \alpha \in [0.999\alpha_4, y_j].$$

In this case we get $y_0 < y_1 < \dots < y_7 \approx 0.381 > \tilde{\alpha}_4$. Therefore $\hat{H}_k(\alpha; \lambda_4) < 0$ for all $\alpha \in [0.999\alpha_4, 1/2]$, whence

$$H_4(\alpha, \hat{\mathbf{z}}(\alpha; c); c) < 0, \quad \forall \alpha \in (0, 1/2], \quad c < 1.$$

The reader may wonder why we did not use a cruder recurrence

$$y_j = H^{-1} \left(\ln \frac{2}{1 + e^{-2.4y_{j-1}}} \right), \quad (y_0 = 0.999\alpha_4).$$

The reason is that in this case $\lim_{j \rightarrow \infty} y_j \rightarrow \bar{y} \approx 0.274$, the root of

$$H(y) + \ln \frac{1 + e^{-2.4y}}{2} = 0.$$

That is, this sequence $\{y_j\}$ forever stays below $\tilde{\alpha}_4 \approx 0.374$!

Appendix (b). Suppose $\alpha \in (0, 1/2]$. Let us show that $\partial \bar{H}_k(\alpha; \lambda)/\partial \lambda > 0$ if $\bar{H}_k(\alpha; \lambda) \geq 0$. Introduce

$$(54) \quad \begin{aligned} u &= u(\alpha) := e^{-H(\alpha)/k} \left[(1 - \alpha)^{\frac{k-1}{k}} - \alpha^{\frac{k-1}{k}} \right], \\ v &= v(\alpha) := e^{-H(\alpha)/k} \left[(1 - \alpha)^{\frac{k-1}{k}} + \alpha^{\frac{k-1}{k}} \right], \end{aligned}$$

so that

$$\bar{H}_k(\alpha; \lambda) = \frac{\psi(\lambda)}{k} H(\alpha) + \ln \frac{f(u\lambda) + f(v\lambda)}{2f(\lambda)}, \quad \left(\psi(x) := \frac{xf'(x)}{f(x)} \right).$$

Obviously, $u \leq 1$ and it can be shown that $v \geq 1$. The condition $\bar{H}_k(\alpha; \lambda) \geq 0$ and log-convexity of $f(x)$ imply that

$$(55) \quad \frac{\psi(\lambda)}{k} H(\alpha) \geq \ln \frac{2f(\lambda)}{f(u\lambda) + f(v\lambda)} \geq \ln \frac{2}{e^{\psi(\lambda)(u-1)} + e^{\psi(\lambda)(v-1)}}.$$

Further

$$(56) \quad \frac{\partial}{\partial \lambda} \bar{H}_k(\alpha; \lambda) = \frac{1}{k} \psi'(\lambda) H(\alpha) + \frac{\partial}{\partial \lambda} \ln \frac{f(u\lambda) + f(v\lambda)}{2f(\lambda)}.$$

Here

$$\frac{\partial}{\partial \lambda} \ln \frac{f(u\lambda) + f(v\lambda)}{2f(\lambda)} = \frac{\left(\frac{f(u\lambda)}{f(\lambda)} \right)'_{\lambda} + \left(\frac{f(v\lambda)}{f(\lambda)} \right)'_{\lambda}}{\frac{f(v\lambda)}{f(\lambda)} + \frac{f(v\lambda)}{f(\lambda)}},$$

and, using concavity of $\psi(x)$,

$$\begin{aligned} \left(\frac{f(u\lambda)}{f(\lambda)} \right)'_{\lambda} &= \frac{f(u\lambda)}{\lambda f(\lambda)} [\psi(u\lambda) - \psi(\lambda)] \\ &\geq \frac{f(u\lambda)}{\lambda f(\lambda)} \psi'(\lambda)(u\lambda - \lambda) = \frac{f(u\lambda)}{f(\lambda)} \psi'(\lambda)(u - 1), \end{aligned}$$

and likewise

$$\left(\frac{f(v\lambda)}{f(\lambda)} \right)'_{\lambda} \geq \frac{f(v\lambda)}{f(\lambda)} \psi'(\lambda)(v - 1).$$

So, using also log-convexity of $f(x)$ and $u - 1 \leq 0$,

$$\begin{aligned}
 (57) \quad \frac{\partial}{\partial \lambda} \ln \frac{f(u\lambda) + f(v\lambda)}{2f(\lambda)} &\geq \frac{(v-1) + \frac{f(u\lambda)}{f(v\lambda)}(u-1)}{1 + \frac{f(u\lambda)}{f(v\lambda)}} \psi'(\lambda) \\
 &\geq \frac{(v-1) + (u-1) \exp[\psi(v\lambda)(u-v)/v]}{1 + \exp[\psi(v\lambda)(u-v)/v]} \psi'(\lambda) \\
 &\geq \frac{(v-1) + (u-1) \exp[\psi(\lambda)(u-1)]}{1 + \exp[\psi(\lambda)(u-1)]} \psi'(\lambda);
 \end{aligned}$$

we used $v \geq \max\{u, 1\}$. Combining (55)-(57), we see that $\partial \bar{H}_k(\alpha; \lambda)/\partial \lambda > 0$ if

$$g(x, y) := \ln \frac{2}{e^x + e^y} + \frac{y + xe^x}{1 + e^x} > 0;$$

here

$$x = \psi(\lambda)(u-1), \quad y = \psi(\lambda)(v-1).$$

By (54),

$$\begin{aligned}
 y - x &= 2\psi(\lambda)e^{-H(\alpha)/k} \alpha^{\frac{k-1}{k}}, \\
 x &= \psi(\lambda)e^{-H(\alpha)/k} \left[(1-\alpha)^{\frac{k-1}{k}} - \alpha^{\frac{k-1}{k}} - e^{H(\alpha)/k} \right] \\
 &\leq -\psi(\lambda)e^{-H(\alpha)/k} \alpha^{\frac{k-1}{k}},
 \end{aligned}$$

so that

$$(58) \quad x \leq -\frac{y-x}{2}.$$

Since

$$g(x, y) = \ln \frac{2}{1 + e^{-(y-x)}} - \frac{y-x}{1 + e^{-x}},$$

we see that, under condition (58), $g(x, y) > 0$ if

$$\phi(z) := \ln \frac{2}{1 + e^{-z}} - \frac{z}{1 + e^{z/2}} > 0, \quad \forall z > 0.$$

This is indeed so, since $\phi(0) = 0$, and

$$\begin{aligned}
 \phi'(z) &= (1 + e^z)(1 + e^{z/2})^{-2} \sum_{j \geq 3} r_j z^j, \\
 r_j &:= \frac{1}{j!} \left[\frac{j+1}{2^j} + \left(\frac{j}{3} - 1 \right) \left(\frac{3}{2} \right)^n \right] > 0.
 \end{aligned}$$

□

(Boris Pittel) DEPARTMENT OF MATHEMATICS, OHIO STATE UNIVERSITY, COLUMBUS OH 43210, USA

E-mail address: bgp@math.ohio-state.edu

(Gregory B. Sorkin) DEPARTMENT OF MANAGEMENT, LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE, HOUGHTON STREET, LONDON WC2A 2AE

E-mail address: g.b.sorkin@lse.ac.uk